



AFRL-OSR-VA-TR-2013-0216

Ultra - Low - Power Asynchronous Processor and FPGA Design using Straintronics Nanomagnets

Pinaki Maumder

University of Michigan

**May 2013
Final Report**

DISTRIBUTION A: Approved for public release.

**AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND**

REPORT DOCUMENTATION PAGE
*Form Approved
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to the Department of Defense, Executive Services and Communications Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE (DD-MM-YYYY) 07-04-2013	2. REPORT TYPE FINAL REPORT	3. DATES COVERED (From - To) Nov 15, 2011-Nov. 14, 2012		
4. TITLE AND SUBTITLE Ultra - Low - Power Asynchronous Processor and FPGA Design using Straintronics Nanomagnets		5a. CONTRACT NUMBER		
		5b. GRANT NUMBER FA9550-12-1-0033		
		5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Pinaki Maumder		5d. PROJECT NUMBER		
		5e. TASK NUMBER		
		5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Michigan 2269 Hayward Ann Arbor, MI 48109-2121		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AF Office of Scientific Research 875 N. Randolph St, Room 3112 Arlington, VA 22203		10. SPONSOR/MONITOR'S ACRONYM(S)		
		11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-OSR-VA-TR-2013-0216		
12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION A: APPROVED FOR PUBLIC RELEASE				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT This seed and exploratory grant was funded to generate some initial results that are presented in this report. The project was funded by Dr. Robert Coswell in Nov. 2010 in consultation with Dr. Devanand Shenoy to develop a fusion of strain-based spintronics, as well as spin torque transfer devices, and CMOS VLSI technologies. The project was initiated with the hope to start extensive work in non-volatile and ultra-low-power subthreshold and superthreshold VLSI circuits relevant for a wide spectrum of military and space electronic systems. In this project we demonstrated i) an ultra-low power hearing aid speech processor interfacing with a custom designed SRAM to operate fully in sub-threshold regime (Specs: Operating at 1MHz clock frequency; 600 pJ consumption for each FIR operation); ii) a 128 point FFT/IFFT processor in 65nm technology operating in subthreshold regime (Specs: Operates at 1 MHz; energy consumption of 31 nJ/FFT); iii) a sub-threshold operating asynchronous 8051 microcontroller (A8051) with a novel 16T SRAM cell 250 mV. New 16T SRAM block consumes 5.44 pJ for writing and 9.08 pJ for reading) and iv) a 2 KB				
15. SUBJECT TERMS				
16. SECURITY CLASSIFICATION OF: a. REPORT U b. ABSTRACT U c. THIS PAGE U		17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON 19b. TELEPHONE NUMBER (Include area code)

*Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18*

DARPA/AFOSR FINAL PROJECT REPORT

**Title: Ultra - Low - Power Asynchronous Processor and
FPGA Design using Straintronics Nanomagnets**

**DARPA Program Manager: Dr. Devanad Shenoy
DARPA POC: Dr. Robert Coswell, MTO Director
Program Manager: Dr. Gernot Pomrenke, AFOSR**

**Principal Investigator: Professor Pinaki Mazumder
University of Michigan, Ann Arbor, MI 48109
Grant Number: FA9550-12-1-0033
Period: 11/15/2011 - 11/14/2012**

Funded Amount: \$151,157

This seed and exploratory grant was funded to generate some initial results that are presented in this report. The project was funded by Dr. Robert Coswell in Nov. 2010 in consultation with Dr. Devanand Shenoy to develop a fusion of strain-based spintronics, as well as spin torque transfer devices, and CMOS VLSI technologies. The project was initiated with the hope to start extensive work in non-volatile and ultra-low-power subthreshold and superthreshold VLSI circuits relevant for a wide spectrum of military and space electronic systems.

In this project we demonstrated i) an ultra-low power hearing aid speech processor interfacing with a custom designed SRAM to operate fully in sub-threshold regime (Specs: Operating at 1MHz clock frequency; 600 pJ consumption for each FIR operation); ii) a 128 point FFT/IFFT processor in 65nm technology operating in subthreshold regime (Specs: Operates at 1 MHz; energy consumption of 31 nJ/FFT); iii) a sub-threshold operating asynchronous 8051 microcontroller (A8051) with a novel 16T SRAM cell for improved performance and reliability (Specs: Consumes 91.6 nW at 250 mV. New 16T SRAM block consumes 5.44 pJ for writing and 9.08 pJ for reading) and iv) a 2 KB nonvolatile straintronics memory with 1.3 pJ read power.

A follow-up grant is, therefore, requested to support three graduate students who have enthusiastically worked on this project for one year and are now poised to conduct more creative investigations in this promising emerging technology.

1. Ultra-Low Power FIR Channel Bank For Hearing Aid Speech Processor

In this project we demonstrated an ultra-low power hearing aid speech processor interfacing with a custom designed SRAM to operate fully in sub-threshold regime. The overall FIR system consumes $10.4\mu\text{W}$ while operating at 1MHz clock frequency. Each FIR operation takes 60 clock cycles leading to 600pJ consumption for each FIR operation.

1.1. Motivation

Trend of integrated microsystems demands lower area and power for implantable devices. This is due to the fact that these devices have limited energy harvesting ability and the battery cannot be repeatedly changed. Also as CMOS technologies scale down the leakage power starts to dominate the system. Sub-threshold circuit design is one of the remedies to reduce high static and dynamic power dissipation. Fig 1 shows how VDD reduction can assist power saving. This approach is even more effective in low performance applications including biomedical devices since they reach few kilohertz at their best.

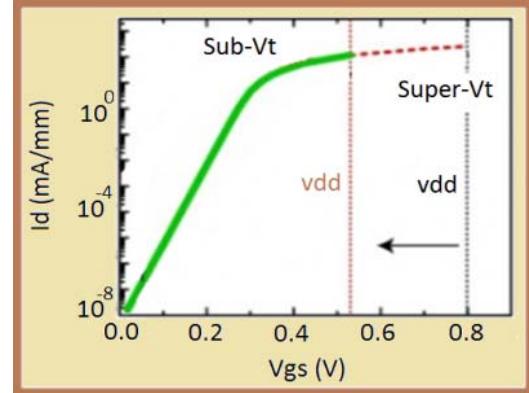


Figure 1. Voltage reduction and its effect on power reduction

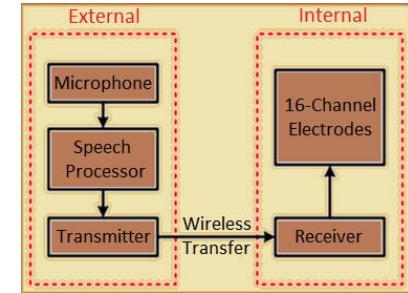


Figure 2. Internal and external parts of cochlear system

1.2. Cochlear implants

They are the oldest and most successful biomedical implant with more than 150,000 worldwide users. They include internal and external part as show in Fig 2. They external part carries a battery and is limited due to its battery life. Speech processor is the power hungry unit of the system. Engineering this unit for power saving can significantly increase the system's battery life.

1.3. FIR filters

FIR filters are typically chosen for the design of cochlear implants since they are inherently stable. They require no feedback since all the poles of the system are located in the middle of the unit circle. They can also be easily designed for linear phase if the impulse response is made symmetrical.

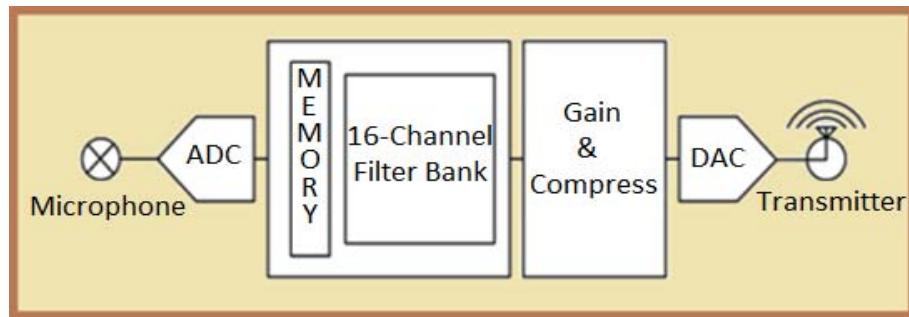


Figure 3. Units of the outside system and demonstration of where the memory and filter bank are placed.

1.4. System Architecture and Memory

The overall unit that is outside of ear is shown in Fig 3. The proposed memory plus FIR system lays between ADC and the gain stage. The block diagram of the proposed design is shown in Fig 4. By choosing the digital design the dynamic power reduces significantly at low frequencies since $P_{dyn} = C \times VDD^2 \times f_{clk}$. The overall function to achieve with the 110 tap FIR system is given by:

$$Y[n] = \sum_{k=0}^M h^j[k] \times \{X[n - k] + X[n - 2M - 1 + k]\}$$

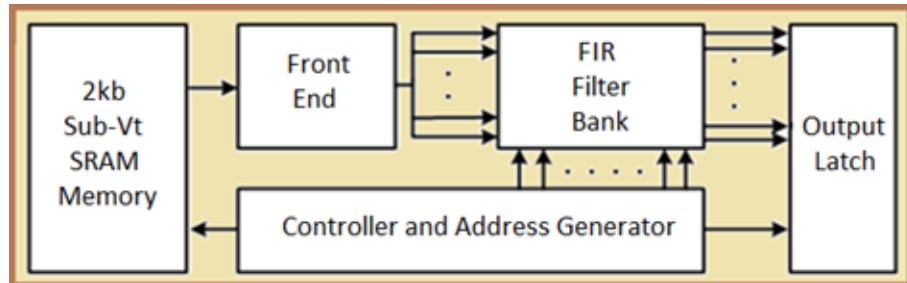


Figure 4. Overall proposed architecture

The channel bank which is shown in Fig 5 follows the ANSI S1.11 1/3 octave standard in which:

$$f_c(j) = 2^{\left(\frac{j-30}{3}\right)} \times f_r$$

i-th center Frequency

$$\Delta f(j) = f_c(j) \times (2^{\frac{1}{6}} - 2^{-\frac{1}{6}})$$

i-th filter BW

The coefficients are generated inside the controller. The architecture avoids multipliers and shares the coefficients generated from the front end block.

The frequency response is a non-uniform as shown in Fig 6 response that matches the human hearing system.

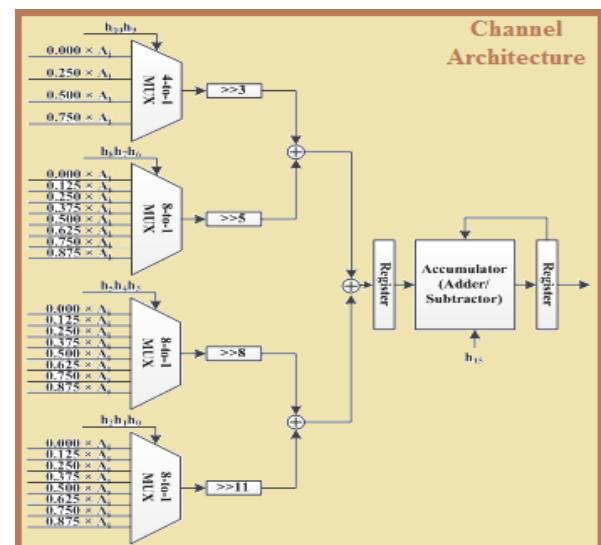


Figure 5. Channel bank architecture

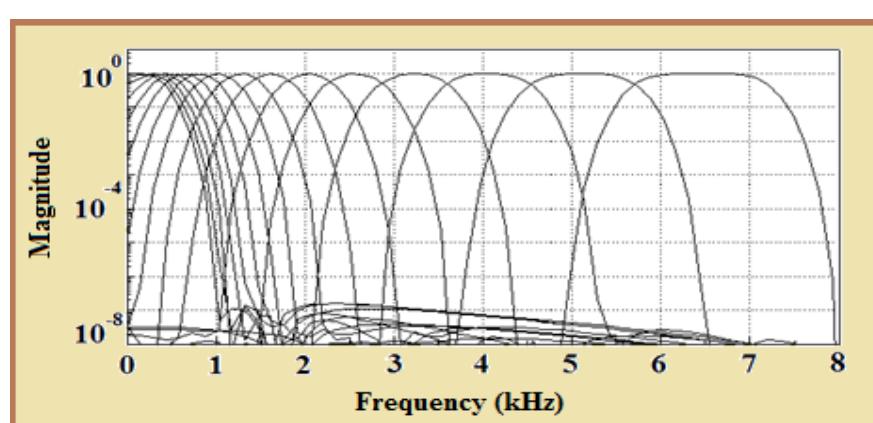


Figure 6. Frequency response of the FIR system

1.4.1. Memory Design and simulations

In order to operate in sub-threshold regime reliably a custom 10T SRAM with dual read is developed. It's higher noise margin and lower Bit Error Rate (BER) is demonstrated in Fig 7.

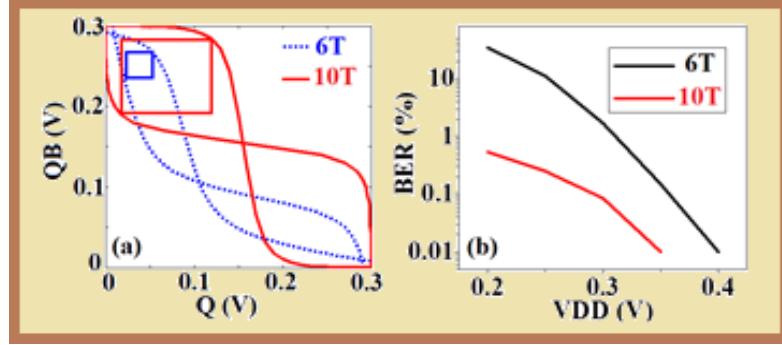


Figure 7. SNM and BER plots comparing the conventional 6T with the proposed 9T cell

The lower leakage of this cell is another advantage which is demonstrated in Fig 8. The conventional 6T shows lower leakage at higher VDD since it has less components but shows higher leakage compared to the 10T at lower voltages. This is due to the shielding of the back to back inverter from the BL that is provided in the 10T cell.

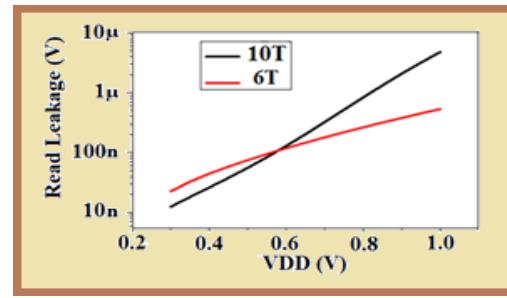


Figure 8. Comparison of read leakage

1.5. Simulation Results and Comparison

The entire system operates at 0.3V in 65nm CMOS technology. Characterization approach is used to reduce the critical path delay at sub-threshold regime. Fig 9 shows the advantage of characterization where 24% higher throughput is achieved with only 1.5% power tradeoff.

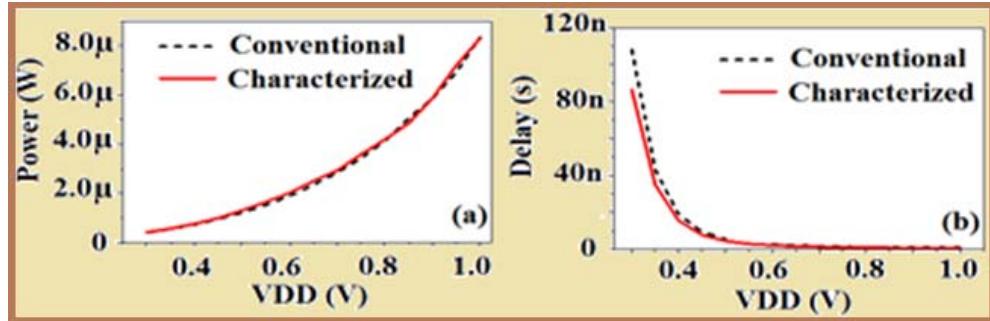


Figure 9. Comparison of read leakage

The power breakdown for the memory and the power breakdown of the entire system at 0.3V supply voltage are shown in Table I and II respectively. The memory consumes almost 32% of the system power thanks to the custom design. This number is much higher for FIR filters with synthesized memories.

TABLE I. Power breakdown of the memory blocks

Total memory power	SRAM bank	Controller and decoders	Buffers and SA's
4.8 μ W	2.1 μ W	1.2 μ W	1.5 μ W

TABLE II. Power breakdown of the entire system

	Front Block	Channel Bank	Cont. & addr seq	Output latch	Memory
Power (μ W)	0.4	7.6	1.3	1.1	4.8
Percentage (%)	2.6	50	8.6	7.2	31.6

The system is simulated for different supply levels and the energy, delay and EDP is shown if Fig 10. The minimum EDP happens at values above sub-threshold. However since we target the low energy dissipation here, we prefer to operate at low voltages to save more energy.

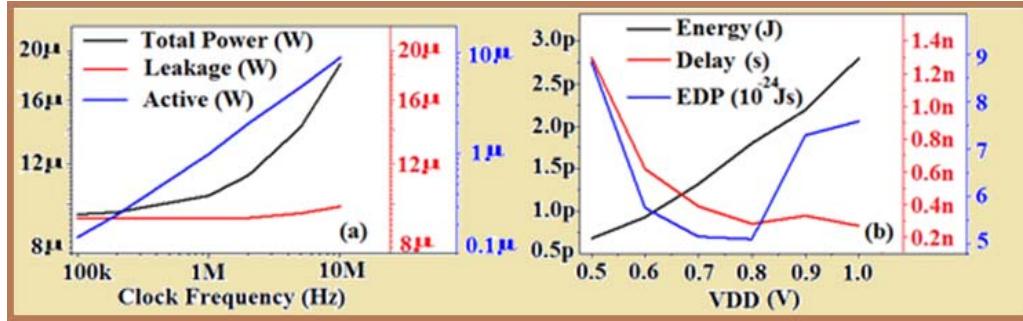


Figure 10. Comparison of read leakage

Due to unit sharing, custom SRAM, and sub-threshold operation the system's energy per FIR operation shows a dramatic reduction compared to the works in the literature. Table III shows the improvements in terms of energy and power compared to the recent works in this field.

TABLE III. Comparison of filter banks in literature

	Technology	Clock frequency	Supply voltage (V)	Power (μ W)	Energy/FIR operation (nJ)
This work	65nm	0.96 MHz	0.3	10.4	0.6
[1]	350nm	1 MHz	1.1	220	12
[2]	130nm	6.13 MHz	0.6/1.2	87	3.6

2. Subthreshold FFT Processor for Low-power Mobile Applications

In this project we have designed a 128 point FFT/IFFT processor in 65nm technology operating in subthreshold regime. The processor incorporates a custom designed subthreshold 4 kb SRAM with 8T unit cells. The subthreshold processor runs at 1 MHz with an energy consumption of 31nJ/FFT.

2.1. Motivation

Fast Fourier Transform (FFT) processing is a key component in integrated applications requiring energy efficiency. For example, it is used in wideband Orthogonal Frequency Division Multiplexing (OFDM) for robust communication systems. FFT reduces power consumption and complexity, and increases transmission bandwidth and efficiency. It is used in wireless sensor networks for tasks such as location sensing, patient monitoring, inventory tracking. These systems are self-contained with finite power sources (i.e. batteries) and can greatly benefit from a low power FFT processors operating in subthreshold regime.

2.2. FFT Computation

Radix-2 FFT computation is picked for hardware simplicity since it allows the mapping of FFT computation to series of basic arithmetic (addition and subtraction) operations. It allows the breakdown of 128-point FFT calculation to 7 stages of 2-point FFT calculations. DFT calculation is represented as:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot W_N^{nk} \quad k = 0, 1, \dots, N - 1$$

Where X_k is the k th transformed point of N points and x_n is the n th point in N points. W_N^{nk} represents the corresponding twiddle factor calculated as:

$$W_N^{nk} = e^{-j\frac{2\pi nk}{N}} = \cos\left(\frac{2\pi nk}{N}\right) - j \sin\left(\frac{2\pi nk}{N}\right)$$

Radix 2 FFT calculation breaks down an N point FFT into $N/2$ point FFTs which enables the mapping of 128-point calculation to 2-point FFT calculations through 7 stages:

$$X_k = \sum_{m=0}^{\frac{N}{2}-1} x_{2m} \cdot W_{\frac{N}{2}}^{mk} + W_N^k \sum_{m=0}^{\frac{N}{2}-1} x_{2m+1} \cdot W_{\frac{N}{2}}^{mk}$$

2.3. FFT Processor

FFT processing involves data intensive computation. One adder, one subtractor and one runtime configurable adder/subtractor is required to perform radix 2 computation. Shifters are required for overflow detection and a look up table is required to store the twiddle factors. 4kbit memory is required to store 128 32-bit data points. Memory system consumes significant

portion of total power due to read and write-backs at the beginning and end of each computation respectively. Overall processor architecture is revealed in Fig 11.

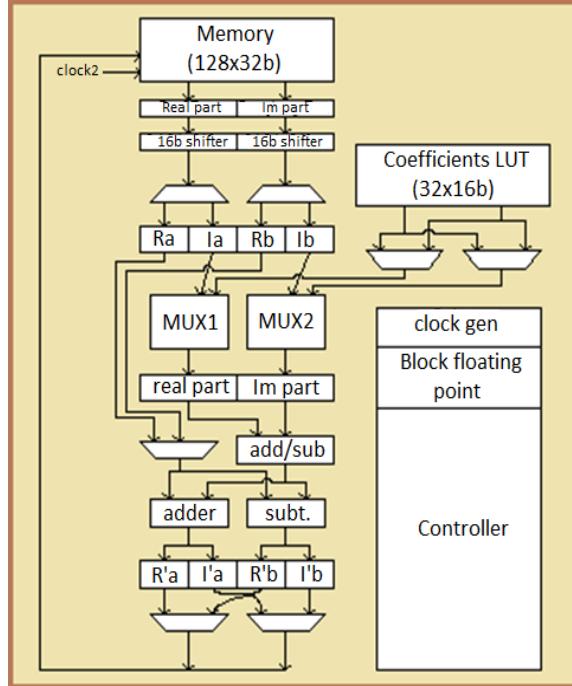


Figure 12 – 128 Point FFT Processor Architecture

2.4. Processor Implementation

The processor is designed to operate in subthreshold operation region to reduce leakage and dynamic power while still meeting the operating frequency requirement of 1MHz. A memory-based architecture (Fig 11) with 5-stage pipeline is used. New data is fetched every two clock cycles to accommodate the proper overlapping (Fig 12) of the pipeline stages.

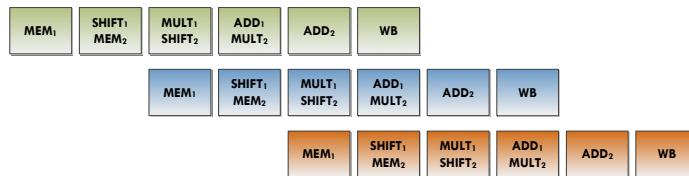


Figure 12 – Pipelined operation stages

Two clock domains (Clock1 and Clock2) are incorporated to increase throughput. The memory is clocked twice as fast as the memory to allow reading from and writing to the memory in the same datapath clock cycle which required the implementation of separate Clock1 and Clock2 domain controllers. Clock1 controller is implemented in a semi-structural fashion which includes an address sequencer as shown in Fig 13. The reason for this is to reduce required hardware compared to a fully behavioral controller.

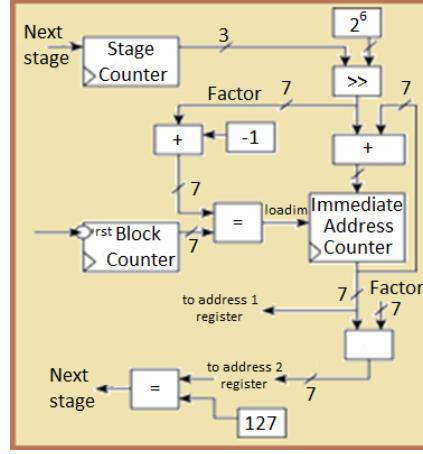


Figure 13 – Semi-structural controller address sequencer

Block Floating Point data representation is adopted in the implementation since it eliminates the use of power hungry floating point units for floating point arithmetic. We have incorporated twiddle factor round off with single bit inaccuracy and ordering to reduce required ROM-storage. The processor operation is fully verified through circuit simulations and achieves >99% accuracy compared to MATLAB FFT/IFFT algorithm.

2.5. Memory

4kbit dual-port 8 transistor SRAM is implemented to operate in subthreshold region. The unit cell structure is shown in Fig 14. Read and write paths are isolated to accommodate for maximum write and read margins allowing aggressive scaling of the supply voltage.

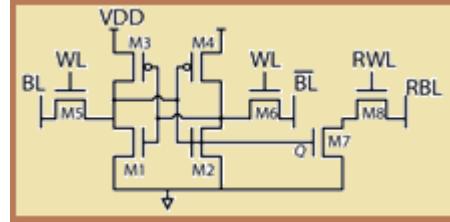


Figure 14 – Unit 8T SRAM cell

The isolation of the read and write paths removes the read margin restrictions on the sizing of the unit cell. The cell design procedure involves resizing the 6T part of the cell and then properly sizing the pass transistors for the 2T reading stack. Therefore, only considering the write margin, the sizing requirement on the transistors becomes:

$$\frac{\beta_5 \cdot I_{OFF}^N \cdot \exp\left(\frac{VDD}{nV_T}\right) + \beta_1 \cdot I_{OFF}^N}{\beta_3 \cdot I_{OFF}^P \cdot \exp\left(\frac{VDD}{nV_T}\right)} > \frac{1 - \exp\left(-\frac{\Delta V}{V_T}\right)}{1 - \exp\left(-\frac{VDD - \Delta V}{V_T}\right)}$$

Where β_X corresponds to the dimensions of the X transistor, I_{OFF}^N and I_{OFF}^P correspond to the N transistor and P transistor zero bias currents, n is the ideality factor of the transistor (assuming

N and P type transistors have similar ideality factors), and V_T corresponds to the thermal voltage which is about 26 mV at 300 K.

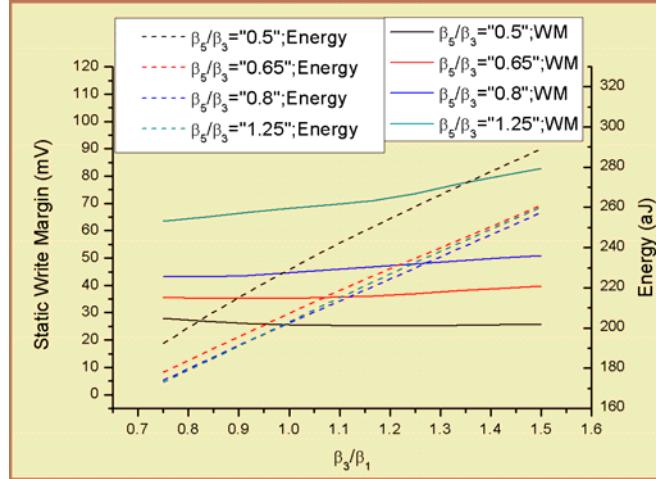


Figure 15 – 8T SRAM cell static write margin as a function of transistor sizing showing inverter sizing is a weaker function of write margin than access transistor sizing

We have conducted studies of transistor sizing as a function of write noise margin and power consumption. The energy study performed looks at the switching and leakage energy for a given time where word line is activated high. Fig 15 also shows that improved write margin is not a strong function of the cross-coupled inverter sizing so in order to reduce leakage in the memory, the ratio of β_3 to β_1 should be kept small.

2.6. Simulation Results

In order to verify the accuracy and correctness of the FFT computation performed by the hardware we have used the signal expressed as:

$$0.5 \times \left(\sin\left(\frac{2\pi n}{50} + \frac{\pi}{4}\right) + j \times \sin\left(\frac{2\pi n}{10}\right) \right)$$

Fig 16 shows the reconstructed signals for both MATLAB and our processor output after the signal is applied an IFF transform following an FF transform.

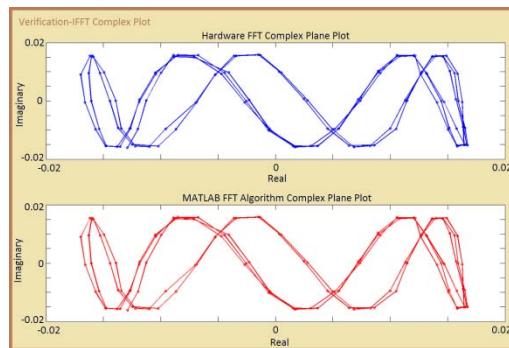


Figure 16 – Reconstructed signals. Upper graph: hardware implementation results. Lower graph: MATLAB results

The input signal is applied FFT using both MATLAB and our processor and the transformed signals and the difference between them (Error) are plotted in Fig 17.

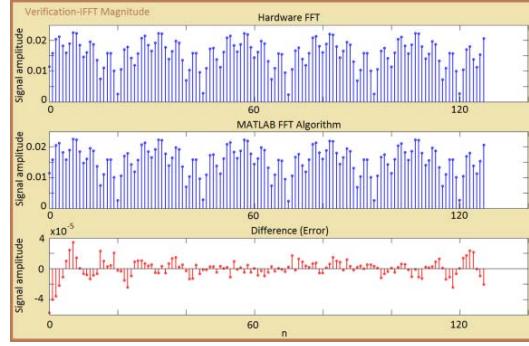


Figure 17 – Graph 1: Hardware FFT Computation magnitude spectrum. Graph 2: MATLAB FFT computation magnitude spectrum. Graph 3: Magnitude difference (error magnitude)

The results indicate that hardware implementation tracks MATLAB to within 0.11%. The small error is due to truncation because the hardware implementation is limited to 16 bits. The error magnitude is dependent on how the initial data is formatted and on the BFP number of shifts to prevent overflow.

Simulations are carried out with 0.3V supply voltage. 65nm technology stand cells are characterized for the subthreshold supply voltage level to obtain accurate critical path delay and optimize accordingly which yielded about 8% delay improvement. With 1MHz clock frequency, FFT processor consumes about 24nJ/FFT and the memory consumes about 7nJ/FFT for a total of 31nJ/FFT energy consumption. Table IV shows comparisons between different power saving techniques for 128 point FFT Processor. The compared results include a synchronous and asynchronous processor operating in superthreshold regime and the processor designed in this work.

TABLE IV. Energy comparison between 128-point FFT processors with different power saving techniques

	VDD (V)	L _{eff} (um)	Energy/FFT (nJ)
Synchronous [3]	1.1	0.35	190
Asynchronous [3]	1.1	0.35	120
Sub-V_t	0.3	0.065	31

3. Asynchronous 8051 microcontroller

In this project, we presented a sub-threshold operating asynchronous 8051 microcontroller (A8051) with a novel 16T SRAM cell for improved performance and reliability. The A8051 consumes 8.98 mW at nominal voltage (1.0 V), while it does 91.6 nW at 250 mV. The embedded novel 16T SRAM block consumes 5.44 pJ for writing and 9.08 pJ for reading.

3.1. Motivation

Wireless sensor networks (WSNs) are increasingly ubiquitous, in part, due to their ultra-low power, high reliability operation, and a small form factor. Fig 18 depicts a typical WSN, comprising six main modules: a sensor, a front-end (e.g. an analog-to-digital converter), a microprocessor, a digital signal processor, a wireless transceiver, and a power management unit including a power source. As this WSN is typically designed for a long operational life-span, power is carefully budgeted where pertinent, and it is energized only when required so that the overall average power is typically 10 µW – 100 µW. The microprocessor module with ultra-low power dissipation is highly desirable as it often remains active for parameter monitoring. Among many processors, 8051 core is still a popular processor for ubiquitous computing since its simplicity as well as low architecture overhead. In addition, sub-threshold operation techniques are a good approach for achieving an ultra-low power solution. However, reliability during sub-threshold operation has been an issue caused by timing uncertainty due to PVT variations. In order to resolve this issue, asynchronous approach is proposed since it is always functionally correct even under PVT variations.

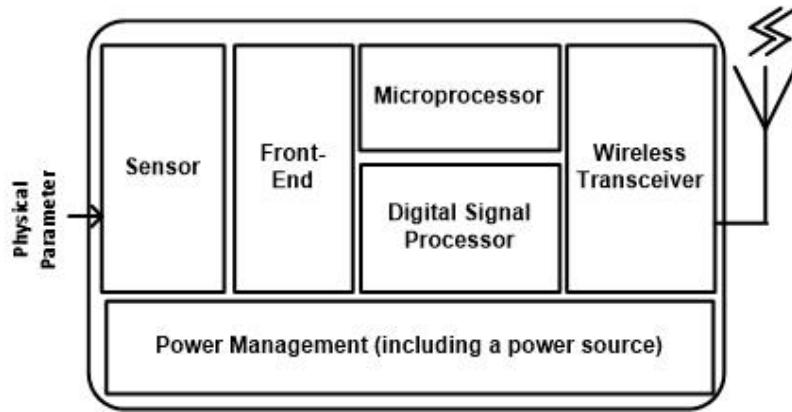


Figure 38. Wireless sensor network

3.2. A8051 Core

A8051 consists of a 1,024×8 read-only memory (ROM) for programming instructions, a 128×8 random-access memory (RAM) for storing data, a 1,024×8 external RAM (XRAM), A8051 core, and an interface block for controlling program ports and general purpose I/O ports. The A8051 core mainly consists of two pipeline stages: Instruction Fetch (IF) and Decode and Execute (D&X) (see Fig 19). IF, Flow Controller (FCont), Instruction Pointer Arithmetic Unit (IPAU), Instruction Pointer (IP), and Memory Controller (MemCont) participate in the first pipeline stage in which instruction fetch and exceptions including interrupts and branching are processed. Instructions are fetched by one byte at each cycle in this IF stage since majority of

instructions have 1-byte length. In the second pipeline stage, operands fetch, operation execution, and write back are conducted by D&X, Arithmetic and Logic Unit (ALU), Register File (ReF), and MemCont, managing operands fetch, executing operation, and writing back. Each pipeline stage is almost independent from each other so that pipeline stalls can be minimized. Although a deeper pipeline stage can increase the overall throughput of A8051, it requires significant area and energy overhead for a pipeline controller to deal with data dependency and control dependency. In addition, pipeline registers might be needed for stalling in the increased pipeline stages, which, in turn, dilutes the strength of asynchronous circuits: no register at all. Notice that each bus between blocks has handshake channels along with data channels so that each block knows when the resulting data should be delivered to the next block.

3.3. SRAM design

Reliability in memory system is one of the most challenging issues since A8051 is targeted at sub-threshold operation. In addition, low power consumption is another design goal. In particular, lower static power consumption is more desirable since this is one of the most benefits in asynchronous systems. To fulfill these two requirements, a novel 16T SRAM cell is proposed. This 16T SRAM cell always operates in static mode during writing and reading even though there is a feedback in a back-to-back inverter structure for holding. The sizing constraint for correct functioning is no issue in this 16T SRAM cell due to the static operation during writing and reading.

3.4. SRAM cell design

Fig 20 shows the proposed SRAM cell. The target operating voltage is 200 mV. Device M1 to M4 are back-to-back inverters, and additional four PMOS devices (M5, M6, M11, and M12) are attached, two of which are connected to the source node of each PMOS of the inverters in parallel. At each output node of each inverter, two additional NMOS devices (M7 to M10) are connected to the ground node in series. These attached eight devices (M5 to M12) are for eliminating contentions between devices during WRITE operation. Accordingly, the new cell configuration always operates in static mode during WRITE operation. In addition, two additional NMOS devices are connected to each storage node for de-coupling the output node from the storage node as in the case of conventional 8T SRAM cells. In particular, one of these

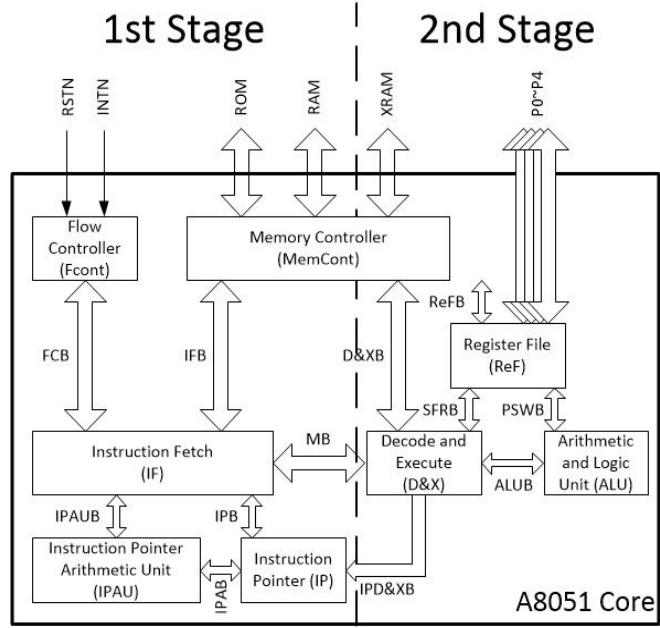


Figure 19. Block diagram of the A8051 core.

transistors is connected to virtual power rail for reducing leakage current to the output bit lines.

3.5. SRAM cell HOLD SNM

Figure 21 shows the corner simulation results for SRAM cell HOLD SNM. These simulations were conducted at $V_{dd}=200$ mV. The best result is 66.6 mV at typical case, while the worst result is 46.3 mV at SF case: slow NMOS and fast PMOS. Total 385,101 transistors are used for the 1,024×8bit ROM, and its cell size is $1.54 \times 1.82 \mu\text{m}^2$ with 40nm CMOS technology.

3.6. SRAM comparison

Table V shows comparison with state-of-the-art sub-threshold SRAM design. The density of this work is the lowest due to the architecture of A8051. The number of transistors is also higher than most previous works. Layout cell size is $1750 \lambda^2$, which is almost twice as conventional 8T SRAM. It can operate at 250mV under process variations, but, in typical, it can even operate at

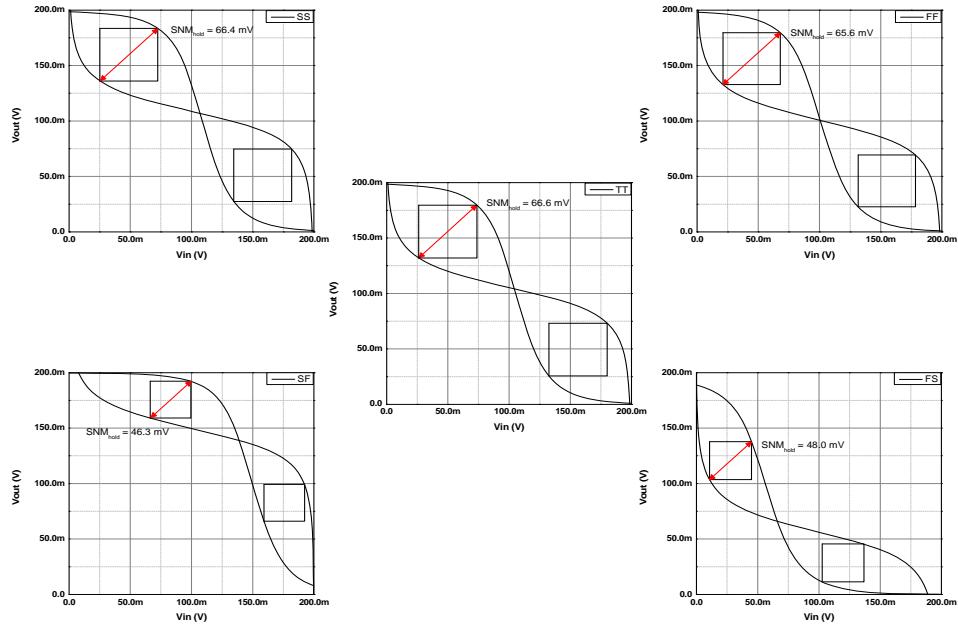


Figure 21. Corner simulation results for SRAM cell HOLD SNM. Best case is at TT: 66.6 mV, and the worst case is at SF: 46.3

200 mV. For writing, performance is relatively high since there is no charge contention during writing. Moreover, power consumption is also lower than the other works because all circuits operate fully in static mode and the density of this work is much lower than the other.

TABLE V. Comparison of state of art SRAM designs

	This work	JSSC'05 [4]	JSSC'07 [5]	JSSC'08 [6]	JSSC'08 [7]	JSSC'09 [8]	JSSC'12 [9]
Technology	40nm	180nm	65nm	65nm	130nm	90nm	65nm
Number of bit lines	128	128	256	256	1k	256	64
Density	1kb	16kb	256kb	256kb	480kb	32kb	72kb
Type	16T	18T	10T	8T	10T	10T	9T
Operation Range	250mV 248.5 kHz 79.3 kHz (Read)	180 mV (164Hz) ~900 mV (6MHz)	0.32V (read) 0.38V (write) 465kHz @ 0.4V	350mV 25 kHz	0.2 V 120 Hz	180mV 31.25Hz	350mV 225Hz
Area (μm^2)	0.53×0.12	2.6×2.1	1.89×1.12	1.89×1.12	4.1×1.5	4×2 (chip)	4x4 (chip)
Area Density ($\mu\text{m}^2/\text{bit}$)	2.80	341.25 (chip)	8.27	8.27	12.81	250.00 (chip)	222.22 (chip)
Normalized Area Density	1.00	121.88 (chip)	2.95	2.95	4.58	89.29 (chip)	79.36 (chip)
Power Consumption	0.34 μW (write) 0.18 μW (read)	115nJ/FF @350mV	3.28 μW	2.2 μW	2.04 μW	1.81 μW (read) 1.07 μW (write)	4.05 μW

3.7. Asynchronous 8051 chip

A8051 chip is fabricated with 40nm CMOS technology. The chip size is $1 \times 1\text{mm}^2$. It consumes 39.57 nW at 200 mV, while 8.63 mW at nominal voltage, which is 1.0 V. Total 795,249 transistors are used: 481,627 transistors for memory blocks and 313,622 transistors for A8051 core and MUXes. Figure 22 shows the layout view of A8051.

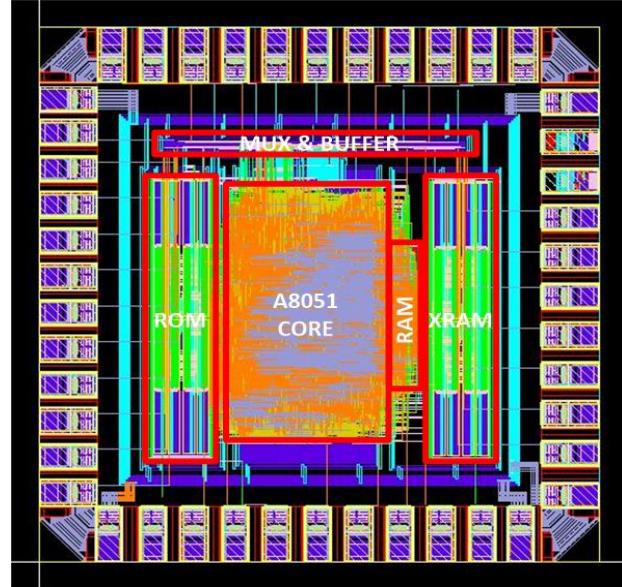


Figure 22. A8051 microcontroller

4. Straintronics Device Modeling

In this project we demonstrated the modeling of the Straintronics device. Then we will interface this model with the CMOS circuitry to develop a nonvolatile memory cell. The memory cells are then combined to form a 2k bit memory. The memory demonstrates much faster read and write speed and better data endurance compared to its flash peer. The latter is due to its magnetic nature. The Straintronics-based memory is low energy and high performance which can be compared to the volatile CMOS SRAM while having the density of a DRAM.

4.1. Motivation

Increasing the power density and leakage to active power ratio as a result of technology downscaling is becoming a concern for CMOS circuit designers. Fig 23 shows the leakage to active power ratio for Intel microprocessors. Volatile CMOS memories like SRAM and DRAM show a high static power consumption to retain the data. However they are the only option to interface with the logic since

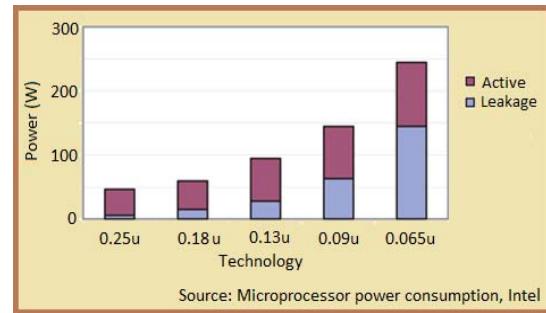


Figure 23. Significant increase of leakage power with technology downscaling

nonvolatile CMOS memories (ex. Flash) are way too slow to fulfill the interface requirements. Besides the fundamental limit for switching of a charge based memory is $NkT\ln(1/p)$ with N being the number of charge carriers. This limit for magnetic logic is $kT\ln(1/p)$ due to magnetic coupling of the domains. However conventional magnetic memories show no power advantage to their CMOS peers since they use current flow for switching. Here we combine piezoelectricity and magnetostriction to obtain a low power switching while keeping speed of the operation. Fig 24 shows where the Straintronics memory (STRRAM) stands among different memory types.

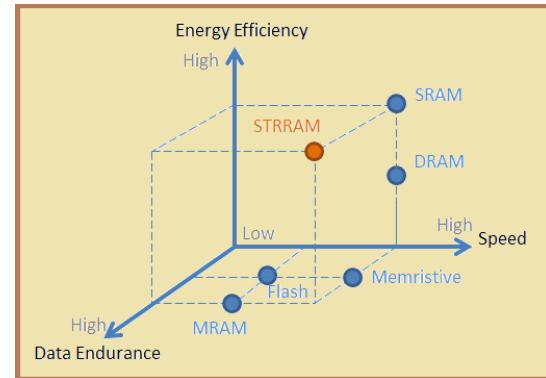


Figure 24. STRRAM lays on <1,1,1> corner of memory design

4.2. Straintronics Principle

Piezoelectricity along with magnetostriction in Fig 25-a leads to the following steps to flip the

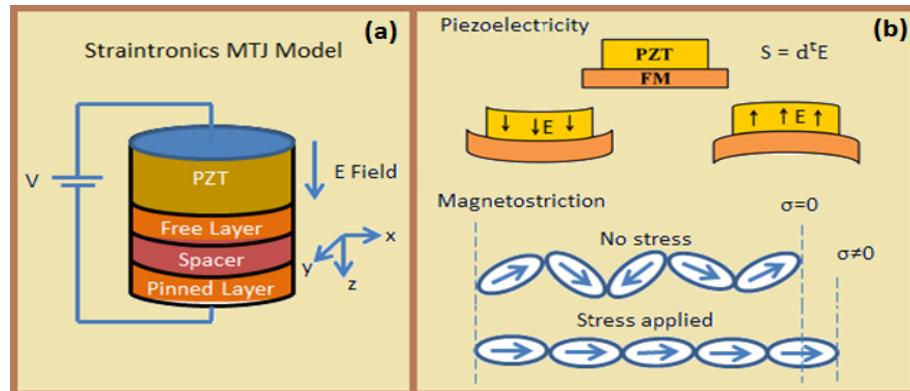


Figure 25. (a) STR device model, (b) piezoelectricity and magnetostriction

magnet. The two phenomena are separately illustrated in Fig 25 –b. i) The E-field causes a strain in PZT leading to a deformation $S = \frac{\Delta L}{L}$. ii) Strain gets transformed to free NM. iii) The magnetostriction effect leads to stress easy axis. iv) At high enough stress, magnet aligns itself

to stress easy axis. As we remove the stress abruptly the magnet continues to flip relaxing at the opposite direction of the starting axis.

4.3. Device Modeling

The STR+MTJ is modeled as an RC circuit. PZT is a parallel place capacitance while MTJ is a variable resistance. This is shown in Fig 26. Shape anisotropy and uniaxial crystalline anisotropy are the primary torques on the magnet when no stress is present. PZT layer is Lead-Zirconate-Titanate with a high dielectric constant. We examined four magnets to find the best fit for the memory cell modeling. Terfenol-D with high magnetostriction coefficient; Nickel with relatively low magnetization saturation; Cobalt with low damping factor; and Metglas 2826MB with high damping factor and low magnetostriction coefficient.

Table VI shows the properties of the four magnets. Also the magnet flipping is demonstrated for the four magnets in Fig 27. Among those, Metglas has a high critical voltage for flipping and therefore provides a high read margin. It also provides a relatively fast response and therefore is chosen for modeling.

3D modeling of the device is based on the LLG equation

Fig 28 shows the flipping of the magnet from P to AP mode and vice versa in the 3D environment. To be clearer we also included the 2D reflection.

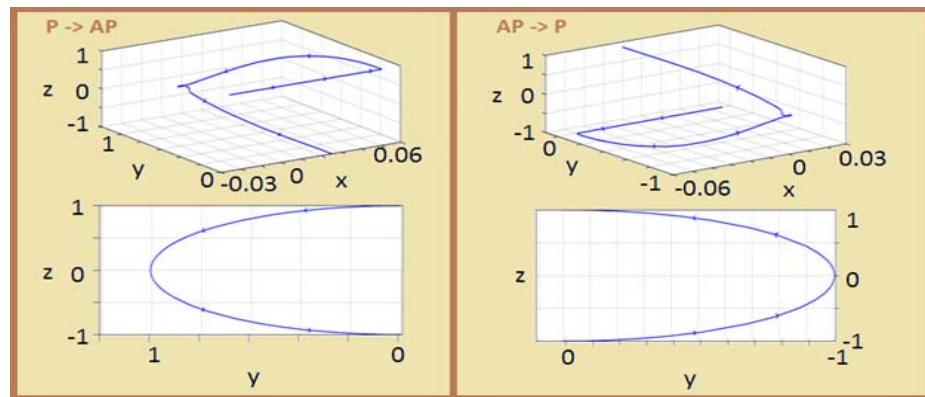


Figure 28. 3D flipping of the magnet

The step by step P to AP flipping of the magnet due to an applied pulse is shown in Fig 29.

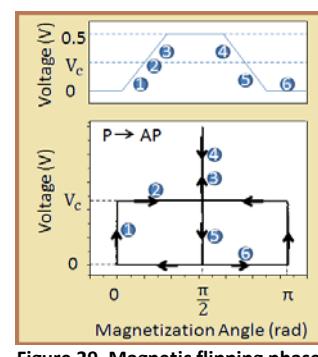


Figure 29. Magnetic flipping phase plot

The stability of the magnet against temperature fluctuations depends on the energy barriers. Here we provide a high barrier ($\sim 80kT$) to prevent unwanted flipping. The voltage applied across the magnet eventually removes the barrier and make the $\theta=90$ the easy axis. This is illustrated in Fig 30-a where the barrier vanishes at 144MPa. The cross section of the flipping is demonstrated in Fig 30-b. The barrier is also a function of the device geometry which is illustrated in Fig 30-c.

4.4. 2kb Memory Design

The proposed memory cell is composed of an STR device

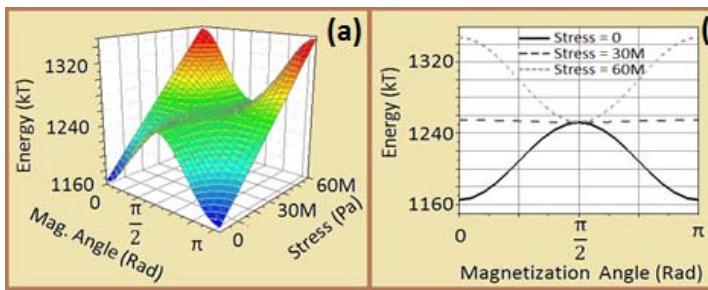


Figure 30. (a,b) Energy barrier as a function of applied voltage (c) Depen

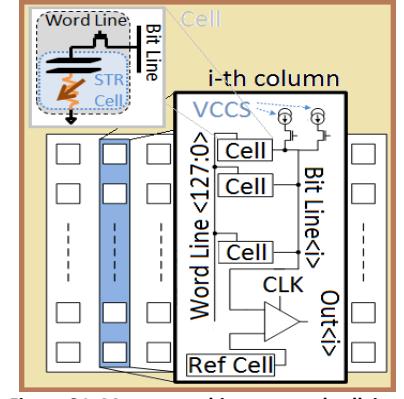


Figure 31. Memory architecture and cellview

and an NMOS access transistor as magnified in Fig 9. The NMOS is chosen to be almost minimum size to maximize the density. Since the MTJ can be laid over the NMOS, the cell can be as small as $0.04\mu\text{m}^2$. The cell is used to build a 2kb memory with 16b simultaneous read/write. The entire memory architecture is demonstrated in Fig 31.

4.4.1. Simulation Results and Comparison

The simulations show merely 1.3pJ read access for 16b access leading to only 80fJ/read access/bit. Write energy is 2.7pJ on average leading to 170fJ/write/bit. Reading can be performed as fast as 355MHz at nominal voltage. Fig 33 shows the read and write energy and read performance of the memory.

The read and write delay as a function of VDD is shown in Fig 34. Due to the critical voltage limitation, we cannot reduce VDD to values lower than 0.8V.

The energy breakdown of the system for read operation is shown in Table VII. To show where the STR memory stands among the memories in the literature, we listed the recent works in literature in table VIII and compared the results to STR memory. STR

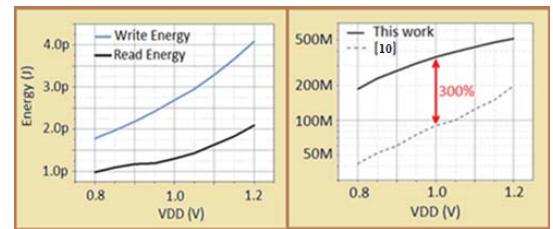


Figure 33. Write/read energy vs. VDD and read performance compared to SRAM in [10]

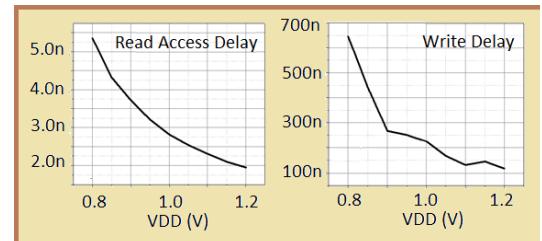


Figure 34. Read/write energy vs. VDD

TABLE VII. Energy breakdown of the system

	Total	Controller	Decoder	Buffers	SA	VCCS
Read Energy (J)	1.3p	319f	52f	117f	258f	554f

memory achieves the energy efficiency of an SRAM, density of a DRAM, durability better than flash and therefore can be a good candidate for a wide range of applications.

TABLE VIII. Energy-performance comparison of STRRAM with memories in recent literature

	Type	Volatility	Magnet /Charge based	Tech	VDD (V)	Energy/Read access (J)	Frequency (Hz)
[10]	SRAM	V	Charge	65nm	0.4	1.4p	475 k
[11]	Flash	NV	Charge	130nm	0.9	76p	50 M
[12]	DRAM	V	Charge	65nm	1	---	500 M
[13]	MRAM	NV	Magnet	90nm	1	900p	66 M
This Work	STR	NV	Magnet	65nm	1	1.3p	355 M

References:

- [1] K. S. Chong; et al, "A 16 Channel Low-Power Nonuniform Spaced Filter Bank Core for Digital Hearing Aids," *Circuits and Systems II: Express Briefs, IEEE Trans. on*, vol.53, no.9, pp.853-857, Sep 2006
- [2] Y. T. Kuo; et al, "Design and Implementation of Low-Power ANSI S1.11 Filter Bank for Digital Hearing Aids," *Circuits and Systems I: Regular Papers, IEEE Trans. on*, vol.57, no.7, pp.1684-1696, July 2010
- [3] K.-S. Chong, B.-H. Gwee, and J. S. Chang, "Energy-Efficient Synchronous-Logic and Asynchronous-Logic FFT/IFFT Processors," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 9, pp. 2034-2045, 2007
- [4] Wang, A. and Chandrakasan, A. P. 2005. A 180-mV subthreshold FFT processor using a minimum energy design methodology. *IEEE Journal of Solid-State Circuits*. 40, 1. 310–319.
- [5] Calhoun, B. H. and Chandrakasan, A. P. 2007. A 256-kb 65-nm Sub-threshold SRAM Design for Ultra-Low-Voltage Operation. *IEEE Journal of Solid-State Circuits*. 42, 3. 680–688.
- [6] Verma, N. and Chandrakasan, A. P. 2008. A 256 kb 65 nm 8T Subthreshold SRAM Employing Sense-Amplifier Redundancy. *IEEE Journal of Solid-State Circuits*. 43, 1. 141–149.
- [7] Kim, T.-H., Liu, J., Keane, J., and Kim, C. H. 2008. A 0.2 V, 480 kb Subthreshold SRAM With 1 k Cells Per Bitline for Ultra-Low-Voltage Computing. *IEEE Journal of Solid-State Circuits*. 43, 2. 518–529.
- [8] Chang, I. J., Kim, J.-J., Park, S. P., and Roy, K. 2011. A 32kb 10T Subthreshold SRAM Array with Bit- Interleaving and Differential Read Scheme in 90nm CMOS. *IEEE Journal of Solid-State Circuits*. 44, 2. 388–390.

[9] Tu, M.-H., Lin, J.-Y., Tsai, M.-C., Lu, C.-Y., Lin, Y.-J., Wang, M.-H., Huang, H.-S., Lee, K.-D., Shih, W.-C. W., Jou, S.-J., and Chuang, C.-T. 2012. A Single-Ended Disturb-Free 9T Subthreshold SRAM With Cross-Point Data-Aware Write Word-Line Operation Timing Tracing. *IEEE Journal of Solid-State Circuits*. 47, 6. 1469–1482.

[10] Calhoun, B.H.; Chandrakasan, A.P.; , "A 256-kb 65-nm Sub-threshold SRAM Design for Ultra-Low-Voltage Operation," *Solid-State Circuits, IEEE Journal of*, vol.42, no.3, pp.680-688, March 2007

[11] Seo, M.K.; et al, "A 0.9V 66MHz access, 0.13um 8M(256Kx32) local SONOS embedded flash EEPROM," *VLSI Circuits, 2004. Digest of Technical Papers. 2004 Symposium on*, vol., no., pp. 68- 71, 17-19 June 2004

[12] Romanovsky, S.; et al, "A 500MHz Random-Access Embedded 1Mb DRAM Macro in Bulk CMOS," *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*, vol., no., pp.270-612, 3-7 Feb. 2008

[13] Nebashi, R.; et al, "A 90nm 12ns 32Mb 2T1MTJ MRAM," *Solid-State Circuits Conference - Digest of Technical Papers, 2009. ISSCC 2009. IEEE International*, vol., no., pp.462-463,463a, 8-12 Feb. 2009